LA-UR-18-22083
Approved for public release; distribution is unlimited.

Title: Modeling cell activation at ultra-low ligand concentrations

Author(s): Suderman, Ryan T.
Hlavacek, William Scott

Intended for: Report

Issued: 2018-03-12

Modeling cell activation at ultra-low ligand concentrations

Ryan Suderman[1,2] and William S. Hlavacek[1,2]

[1]Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory
[2]Center for Nonlinear Studies, Los Alamos National Laboratory

1 Introduction

Very low (femtomolar) concentrations of isoprenaline, a β_2 adrenergic receptor agonist, appear to activate a significant fraction of HEK293 cells grown in 96-well plates. Here, we develop and analyze a mathematical model based on chemical kinetics to determine whether the observed cell activation can be explained by a simple ligand-receptor interaction. Our analysis indicates that a significant fraction of cells can be activated if engagement of one or two receptors is sufficient to trigger a cellular response.

2 Binding model

2.1 Model definition

In this section we define a model to characterize the kinetics of ligand-receptor binding for a population of cells:

$$\mathrm{L} + \mathrm{R}_i \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \mathrm{B}_i \tag{1}$$

where i is an index denoting a particular cell, L represents free ligand, R represents the unbound receptor, B represents the occupied receptor, and k_f and k_r are the association and dissociation rate constants, respectively. Activation of a cell is taken to be proportional to the number of occupied receptors:

$$\mathrm{B}_i + \mathrm{C}_i \xrightarrow{k_{act}} \mathrm{B}_i + \mathrm{C}_i^* \tag{2}$$

where k_{act} is the activation rate constant, C represents an inactive cell and C^* represents an active cell. Note that C_i has a value of 1 until activation and 0 thereafter. In addition to the kinetic parameters, we also introduce f_c, the fraction of cells competent to be activated by the ligand. This parameter is introduced to account for any intracellular conditions (gene expression levels, cell cycle state, etc.) that may prevent a cell from responding to ligand.

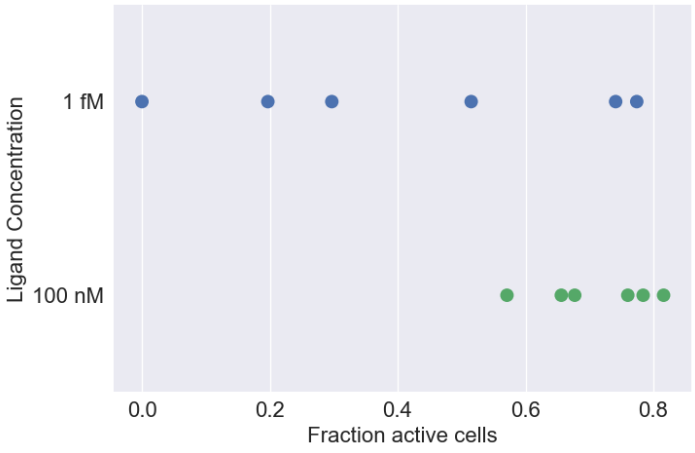

Figure 1: Fraction of responsive HEK293 cells at two distinct doses after a 5 minute exposure to isoprenaline. Prior to ligand stimulation, cells were grown to 70% confluence in 96-well plates. Estimated parameters for the environment are listed in Table 1 and described in the Materials and Methods section in the main text.

2.2 Simulation

For the low ligand concentration of 1 fM, we simulated the stochastic ligand-receptor binding kinetics using Gillespie's algorithm (1). This approach is not computationally feasible when considering the high ligand concentration of 100 nM, because the number of reaction events per unit time scales linearly with the number of molecules in the system. There are 120,440 and 1.2×10^{13} ligand molecules per well for the low and high concentrations, respectively. Note that copy numbers are used for abundances in these simulations, and so the concentrations of the biochemical species and the association rate constant, k_f, must be converted to the appropriate units:

$$\#M = [M] \cdot N_A \cdot V \tag{3}$$

$$k_{f,\#} = \frac{k_f}{N_A \cdot V} \tag{4}$$

where V is the extracellular volume (see Table 1), M is some biochemical species and N_A is the Avogadro constant.

When considering high ligand concentrations (e.g. 100 nM), we make a quasi steady-state approximation for the ligand-receptor interaction. We can then calculate the concentration of occupied receptors in the well since the total concentration of ligand, $[\mathrm{L_T}]$, greatly exceeds the total concentration of receptor, $[\mathrm{R_T}]$:

$$[\mathrm{B}] = [\mathrm{R_T}] \cdot \frac{k_f \, [\mathrm{L_T}]}{k_r + k_f \, [\mathrm{L_T}]} \tag{5}$$

We can also calculate the average concentration of occupied receptors per cell:

$$\langle [\mathrm{B}_i] \rangle = \frac{[\mathrm{B}]}{N_{\mathrm{cells}}} \tag{6}$$

The fraction of cells, F_A, that are active after a time t is:

$$F_A = 1 - e^{-\lambda \cdot t} \tag{7}$$

where $\lambda = k_{act} \cdot \langle [\mathrm{B}_i] \rangle$ is the average rate of activation for each cell. For $k_{act} > 10^{-4}\ s^{-1}$, all cells are activated in less than a minute when $[\mathrm{B}] \approx [\mathrm{R_T}]$.

3 Parameter estimation

We used a Bayesian approach to estimate the following parameters in our model:

- k_r and k_{act}, which are rate constants in the model defined above with units of s^{-1}
- K_D, which is the equilibrium dissociation constant in molar units (M) for ligand-receptor binding and can be used to calculate k_f, given k_r
- f_c, which is the fraction of cells competent for activation (dimensionless)

Our procedure uses a Markov chain Monte Carlo (MCMC) algorithm to estimate a probability distribution for each parameter similar to the procedure outlined in (2). In Bayesian statistics, this estimated distribution is called a parameter's *posterior*. For each parameter set sampled during the MCMC run, estimating the posterior requires calculating both the probability of observing the experimental data given a particular set of parameters (called the *likelihood*) and the probability of the parameters given an assumed probability distribution (the parameter's *prior* distribution).

Two parameters' means and standard deviations have already been characterized in the literature, $\log_{10} K_D$ and k_r (3; 4). We assign $\log_{10} K_D$ to have a normal distribution as its prior:

$$P(\log_{10} K_D) = \text{Normal}(\mu = -9.768, \sigma = 0.612) \tag{8}$$

where μ and σ are the mean and standard deviation of a normal distribution, respectively. Assuming normality for k_r results in significant probability density for values below zero. We therefore assign k_r to have a gamma distribution as its prior, where the gamma distribution's parameters (α and β) are calculated such that the distribution's mean (α/β) and standard deviation ($\sqrt{\alpha/\beta^2}$) correspond to the mean and standard deviation reported in the literature, 0.05 and 0.0255^2, respectively:

$$\frac{\alpha}{\beta} = 0.05 \tag{9}$$

$$\frac{\alpha}{\beta^2} = 0.0255^2 \tag{10}$$

$$P(k_r) = \text{Gamma}(\alpha = 3.845, \beta = 76.894) \tag{11}$$

The prior for the fraction of competent cells, f_c, can be specified based on our data as follows. We assume that 100 nM ligand is a saturating dose that should activate all competent cells, and so we calculate the mean and

Parameter	Value
Volume of medium	200 μL per well
Number of cells	30,000 per well
Receptors	18,000 per cell

Table 1: Fixed model parameters

standard deviation of the 100 nM data in Fig. 1 and assign f_c to have the normal distribution:

$$P(f_c) = \text{Normal}(\mu = 0.711, \sigma = 0.092) \tag{12}$$

where μ and σ are calculated from the data in Fig. 1. The rate of receptor-dependent cell activation relies on incomplete knowledge of the relevant signaling pathways. However we can still constrain this parameter to have a uniform prior distribution over a finite range. We assume that the activation rate must be sufficiently fast to activate cells given potential values of k_r, and that excessively fast activation rates are not physically realizable. Thus, we set:

$$P(\log_{10} k_{act}) = \text{Uniform}(-4, 2) \tag{13}$$

Other fixed parameters used in the model are recorded in Table 1.

Our MCMC sampling was performed for 1,000,000 iterations with a constant jump size of 0.2 (in log space), and we discarded the first 10,000 points as the burn-in period. Parameter updates were accepted using the Metropolis-Hastings criterion, with approximately 37% of the attempted updates being rejected. The sampling trace for $\log_{10} K_D$ can be seen in Fig. 2 and appears to have reached stationarity. From such a sampling trace, we can characterize the posterior distribution of each parameter. As can be seen in Fig. 3, the posteriors for three of the four free parameters strongly reflect their priors. The exception, k_{act} reveals a posterior that is shifted towards larger values, with near uniformity for parameters larger than 0.01. From the MCMC sampling, we can also extract the *maximum a posteriori probability*, or MAP, estimate. The MAP estimate is the point in parameter space analogous to a "best-fit" parameter set, and it is equivalent to the mode of the posterior distribution.

We can further characterize the correlations between the free parameters by looking at their pairwise scatter plots (Fig. 4). As it turns out, all pairwise relationships result in a Spearman's rank correlation coefficient, ρ,

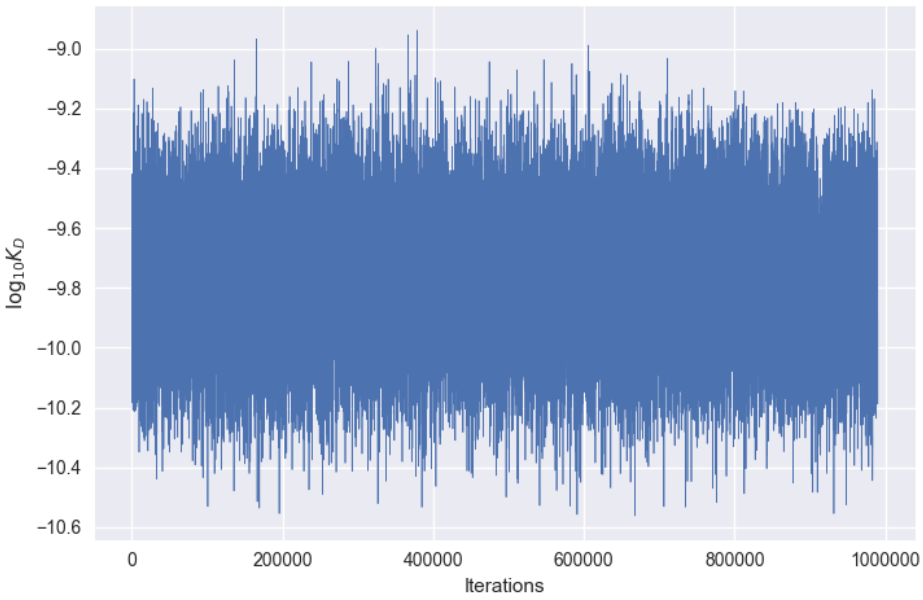

Figure 2: The MCMC sampling trace for $\log_{10} K_D$. The first 10,000 points were discarded as the burn-in period (sampled points prior to reaching stationarity). A total of 990,000 sampled points are shown here.

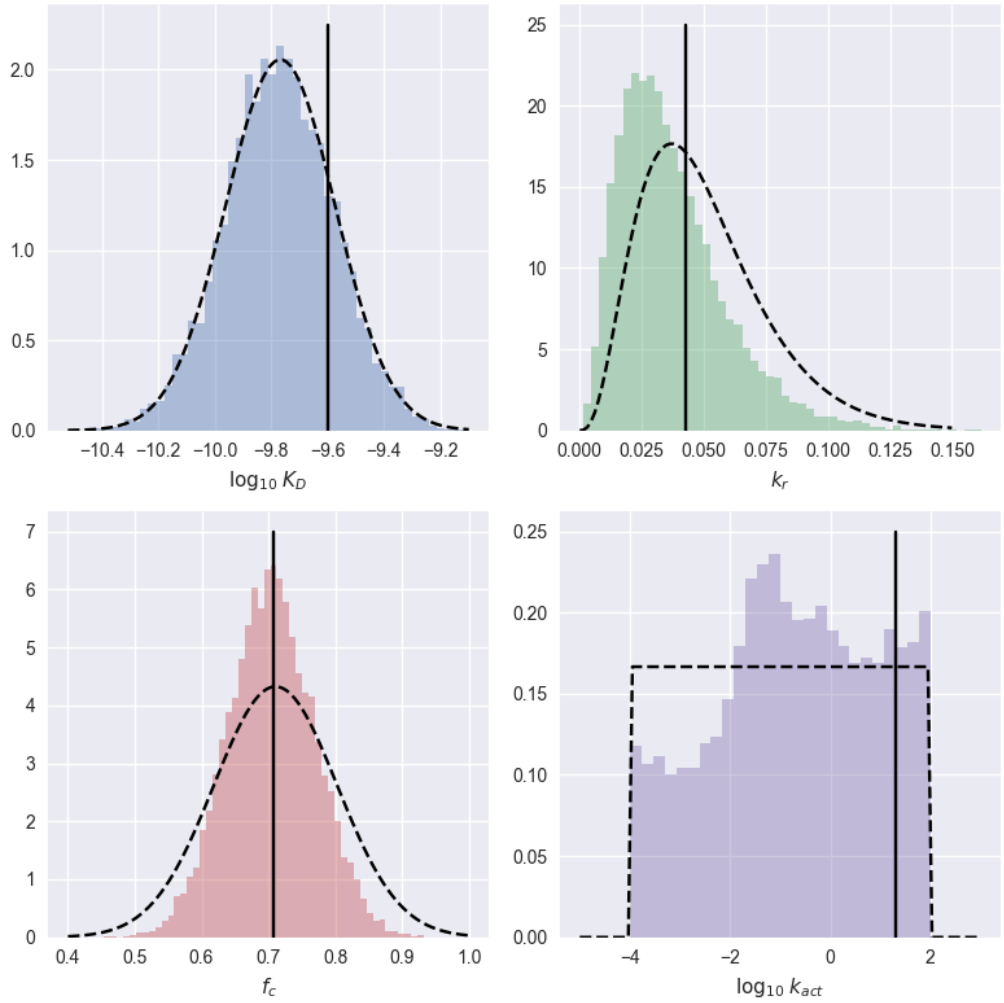

Figure 3: Posterior distributions for the four free parameters. Each black bar indicates the parameter value from the maximum a posteriori probability estimate. Each dashed curve indicates a prior distribution.

of less than 0.05, meaning that dependency between any pair of parameters is unlikely.

4 Model predictions

We can make a number of predictions based on the behavior of the model. Using data from our MCMC sampling approach, we can calculate credible intervals for the timecourse of ligand binding at the low ligand concentration. As can be seen in Fig. 5, the 95% credible interval from 1000 subsampled parameter sets is wide.

We can also characterize the distributions of binding events per cell for the low ligand concentration. Of course this depends on the parameters. For the parameter values corresponding to the maximum a posteriori probability estimate, we found that nearly 70% of the cell population encountered less than 2 ligand molecules, and that about 10% had more than 2 binding events in the allotted time (Fig. 6). Indeed, the average number of binding events is approximately one per cell. The small number of binding events that occur at low ligand concentration are likely reflected in the posterior distribution for k_{act}. For activation to occur in the small number of cells that engage a ligand, the cell must be sufficiently sensitive (i.e., k_{act} must be sufficiently large) to activate the cell within one or two binding events. Our model therefore predicts that the observation of cells responding to femtomolar concentrations of ligand (Fig. 1) requires that the cells must be sufficiently sensitive to respond to just one or two binding events.

Finally, we found the values for the fastest association rate and slowest dissociation rate published in literature ($1.2 \times 10^{10}\,(\mathrm{M} \cdot \mathrm{min})^{-1}$ and $4.8 \times 10^{-4}\,\left(\mathrm{min}^{-1}\right)$, respectively). We inputted these values into our model and fixed k_{act} and f_c to their corresponding values from the MAP estimate parameter set. With this "super ligand" we found that our model predicts an average of one binding event per cell for ligand concentrations as low as 25 attomolar (i.e. 2.5×10^{-17} M)

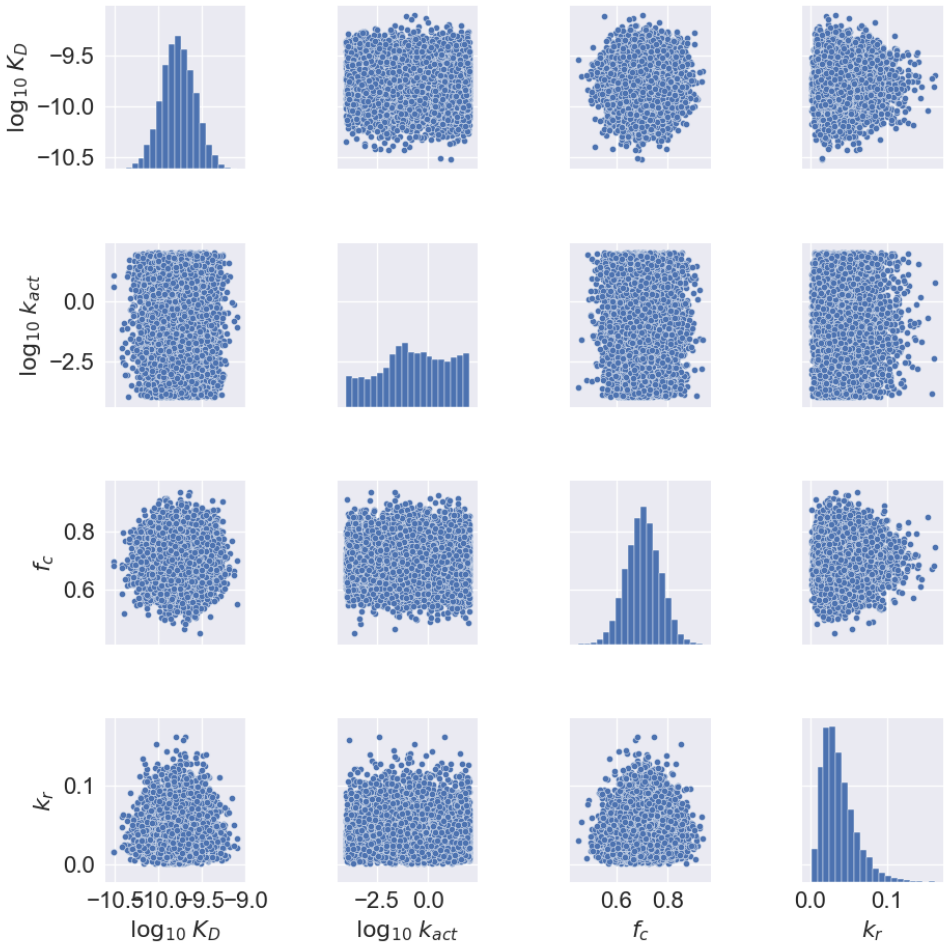

Figure 4: Pairwise relationships between the four free parameters, where the number of points is downsampled for clarity. The diagonal contains downsampled posterior distributions similar to those seen in Fig. 3.

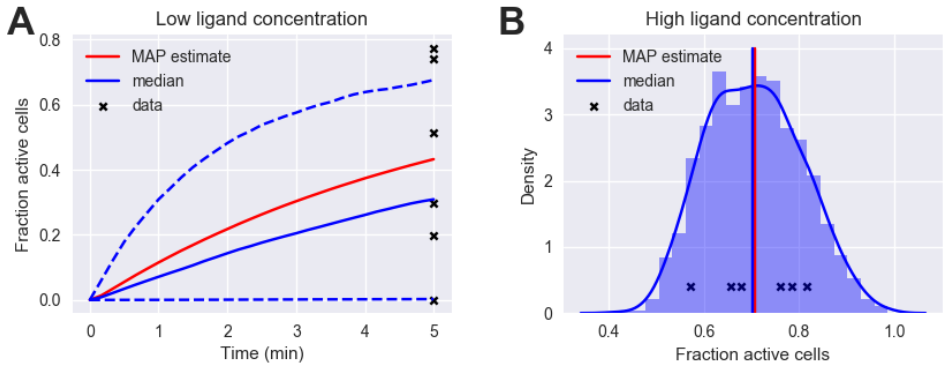

Figure 5: (**A**) The 95% credible interval for 1 fM ligand concentration timecourses using 1,000 randomly subsampled parameter sets from the MCMC run. Shown in red is the timecourse whose parameters correspond to the maximum a posteriori probability (MAP) estimate. The blue solid line corresponds to the median of the subsampled parameter sets, and the blue dashed lines in the timecourse denote the 95% credible interval for the subsampled parameter sets. Note that two of the data points are outliers, resulting from the fact that only a small region ($\sim$ 2%) of sampled parameter space allows the model to reach those points. (**B**) The fraction of active cells for the same subsampled parameter sets, but with the 100 nM ligand concentration. Note that simulations with 100 nM ligand are based on a quasi steady-state approximation. Colors are the same as in panel (A) and black data points in both panels correspond to those in Fig. 1.

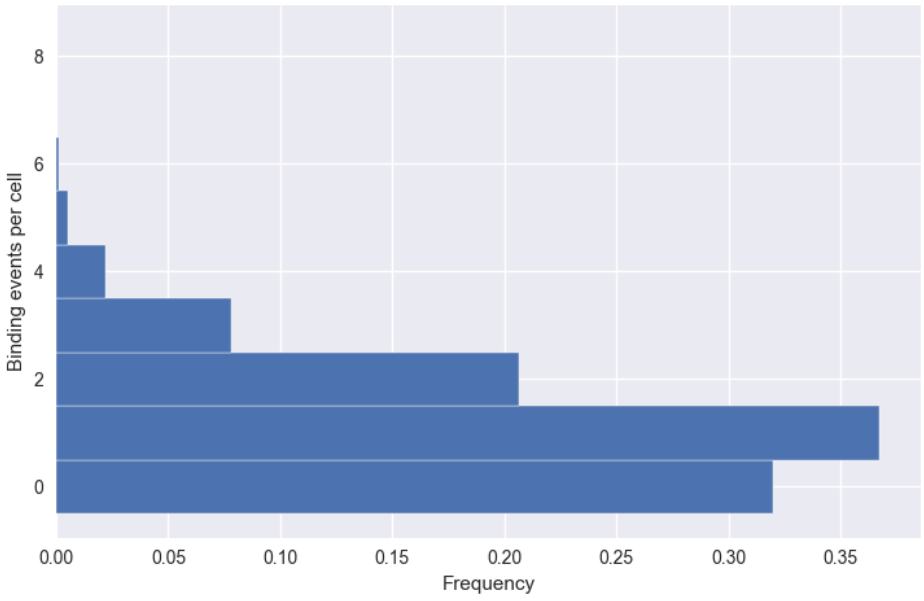

Figure 6: Normalized frequency of binding for 1 fM ligand concentration from 100 independent model simulations with the parameter set corresponding to the maximum a posteriori probability estimate. The average number of binding events is 1.13.

References

1. D. T. Gillespie. Stochastic simulation of chemical kinetics. *Ann. Rev. Phys. Chem.* **58**, 35–55, 2007.

2. N. Kozer, D. Barua, S. Orchard, E. C. Nice, A. W. Burgess, W. S. Hlavacek, and A. H. A. Clayton. Exploring higher-order EGFR oligomerisation and phosphorylation – a combined experimental and theoretical approach. *Mol. BioSyst.* **9**, 1849–1863, 2013.

3. D. P. Staus, R. T. Strachan, A. Manglik, B. Pani, A. W. Kahsai, T. H. Kim, L. M. Wingler, S. Ahn, A. Chatterjee, A. Masoudi, A. C. Kruse, E. Pardon, J. Steyaert, W. I. Weis, R. S. Prosser, B. K. Kobilka, T. Costa, and R. J. Lefkowitz. Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535**, 448–452, 2016.

4. D. A. Sykes and S. J. Charlton. Slow receptor dissociation is not a key factor in the duration of action of inhaled long-acting β2-adrenoceptor agonists. *Brit. J. Pharmacol.* **165**, 2672–2683, 2012.